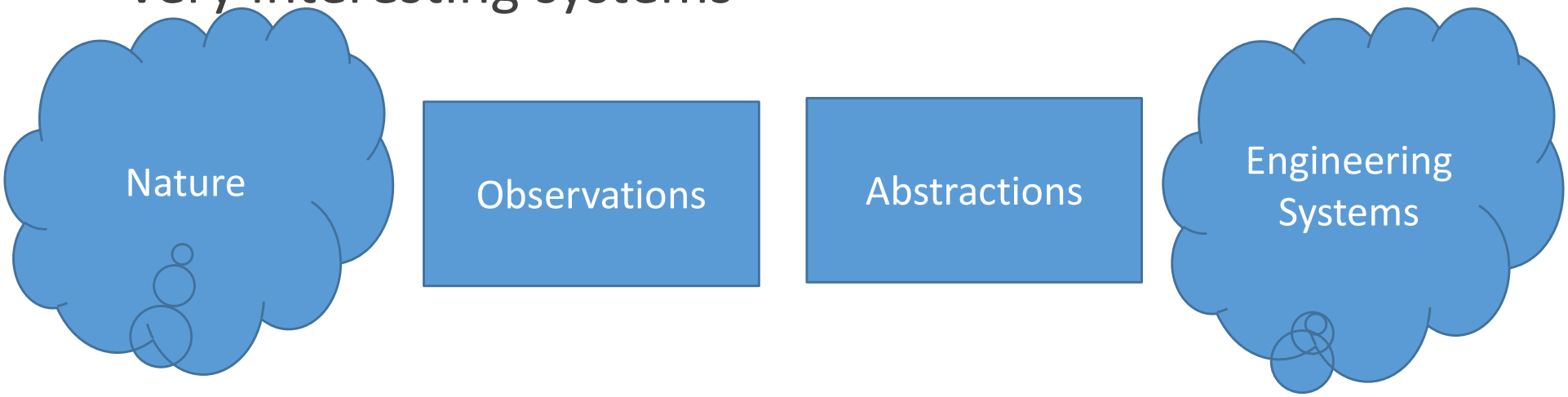
Efficient RISC V Compute Platforms for Enabling the AI Revolution

Dr. Sujay Deb
CTO, AarFive Designs Pvt. Ltd.
Institute Chair Professor, Department of ECE & CSE
Indraprastha Institute of Information Technology Delhi (IIIT Delhi)

Engineering Approach

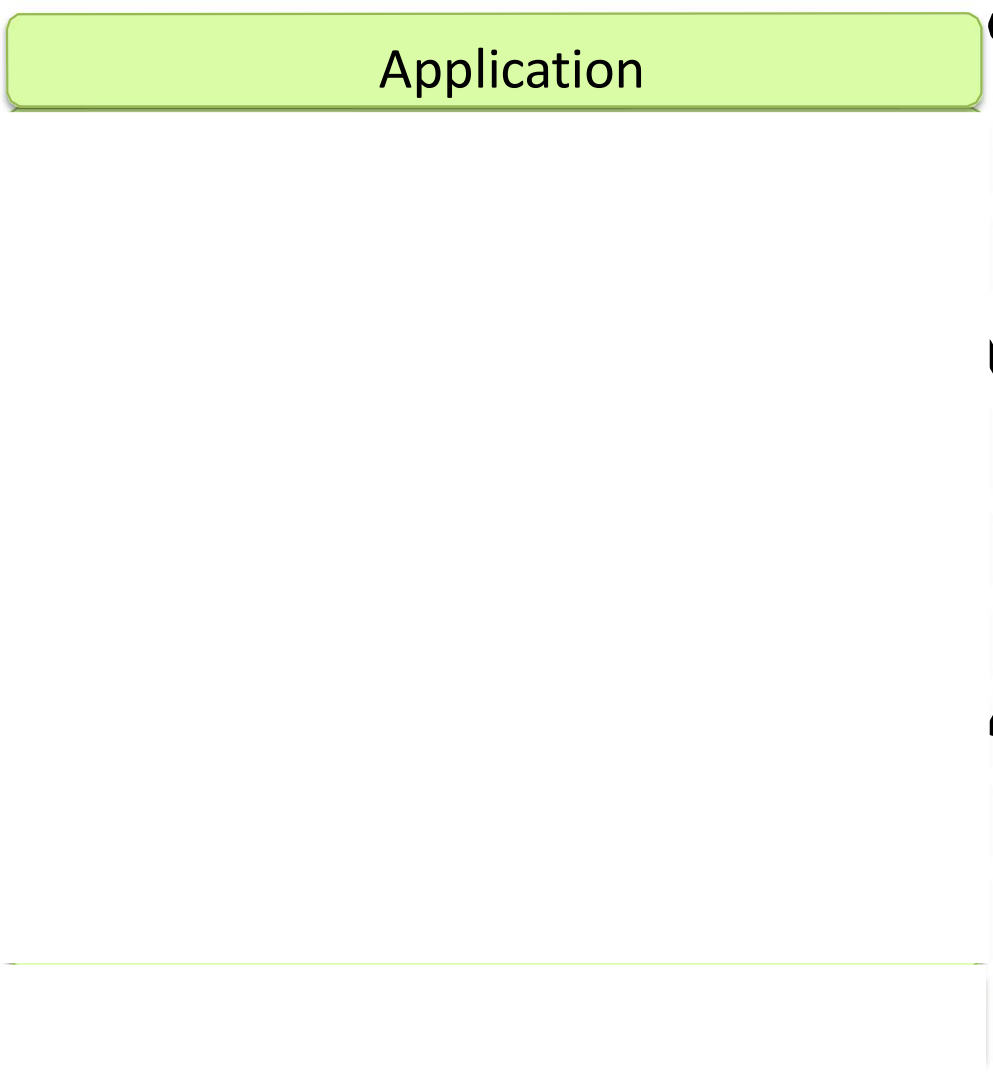
- What is Engineering?
  - the purposeful use of science
- We are here to employ the facts of nature to build very interesting systems



- Take complicated things, build layers of abstraction, and simplify things so that we can build useful systems.

AARFIVE DESIGNS

Abstractions in Modern Computing Systems



Application Requirements:

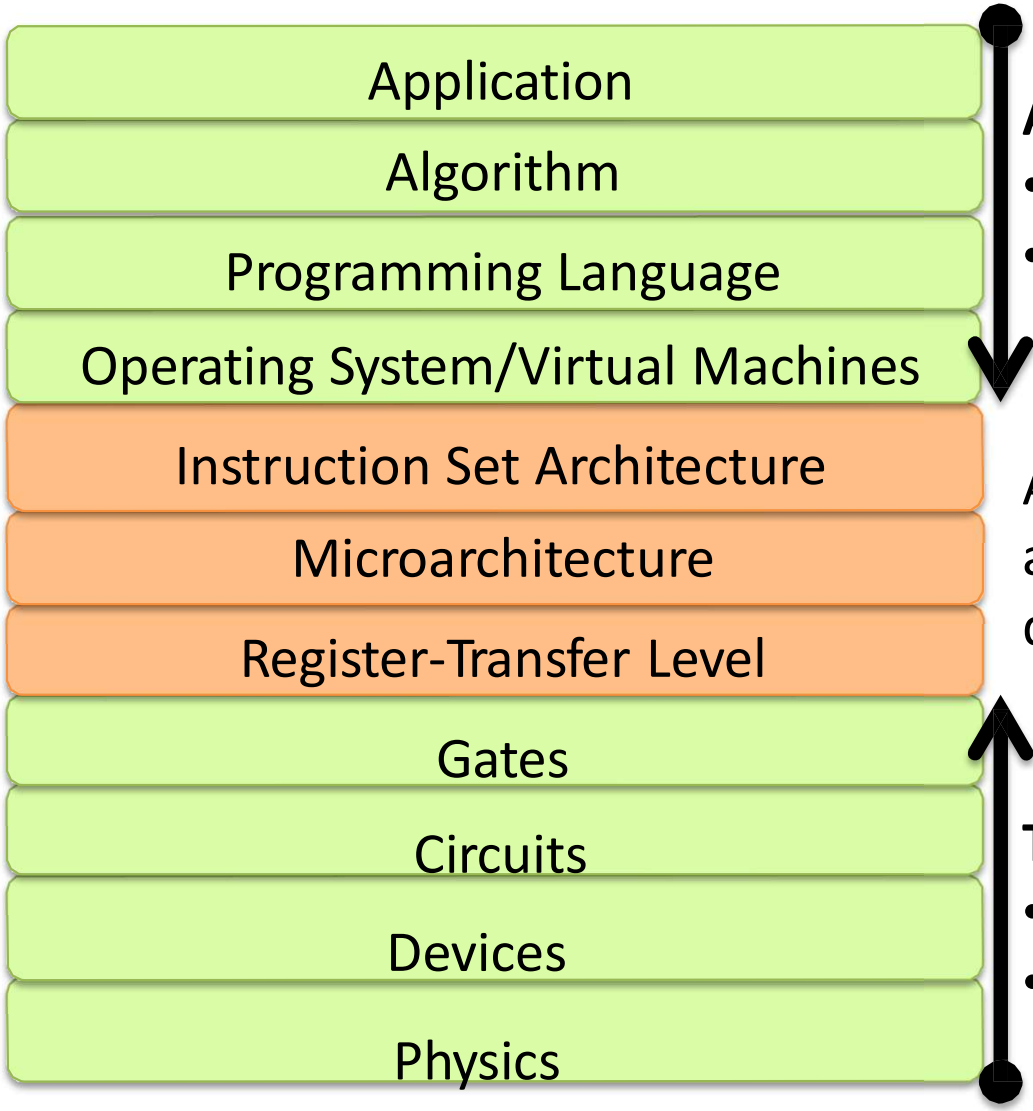
- Suggest how to improve architecture
- Provide revenue to fund development

Technology Constraints:

- Restrict what can be done efficiently
- New technologies make new arch possible

AARFIVE DESIGNS

Abstractions in Modern Computing Systems



Application Requirements:

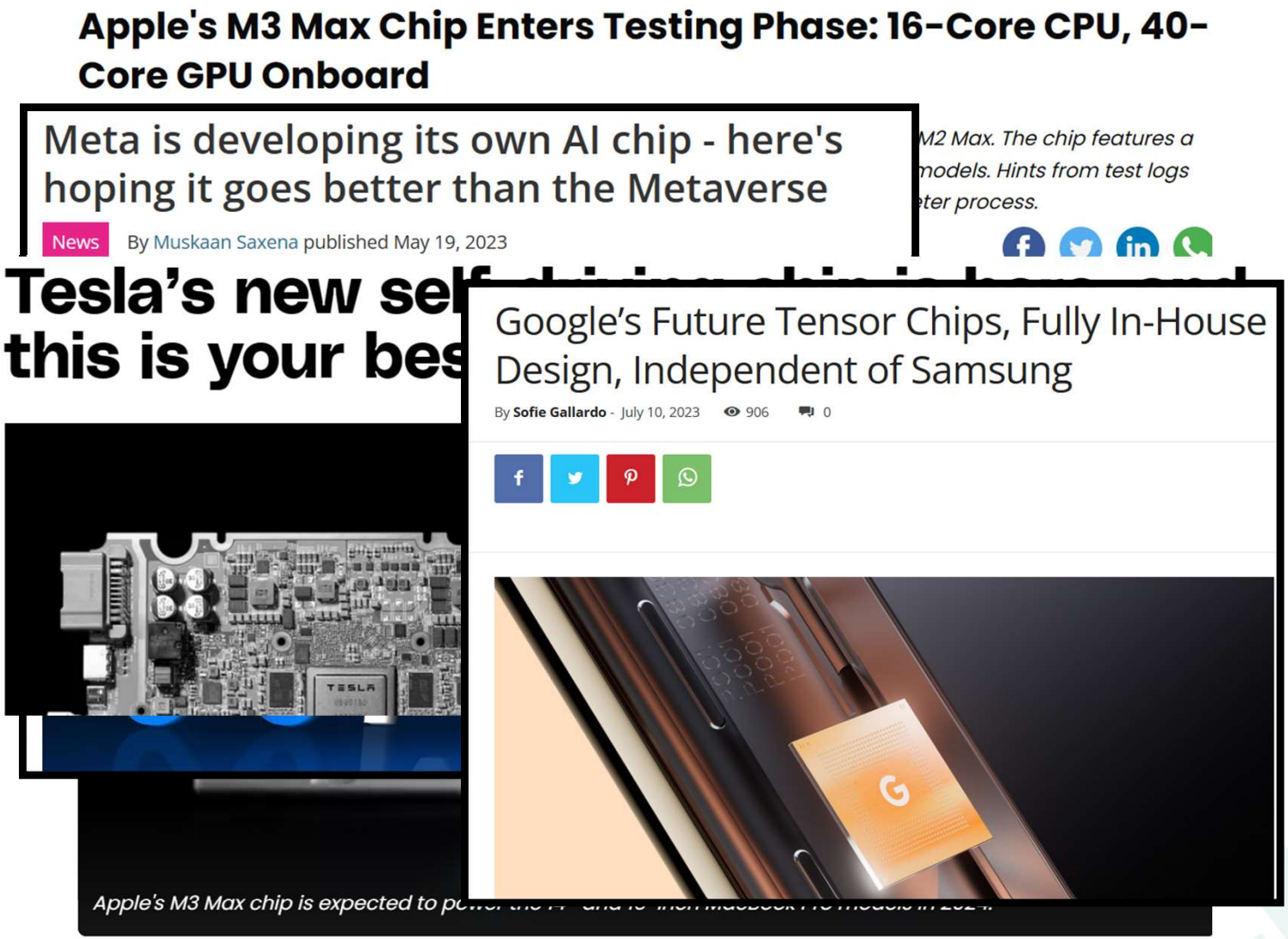
- Suggest how to improve architecture
- Provide revenue to fund development

Architecture provides feedback to guide application and technology research directions

Technology Constraints:

- Restrict what can be done efficiently
- New technologies make new arch possible

AARFIVE DESIGNS



AARFIVE DESIGNS

CPU, APU & CHIPSETS

NVIDIA and MediaTek's new Arm-based AI chip rumored for 1H 2025 to fight Intel, AMD, Qualcomm

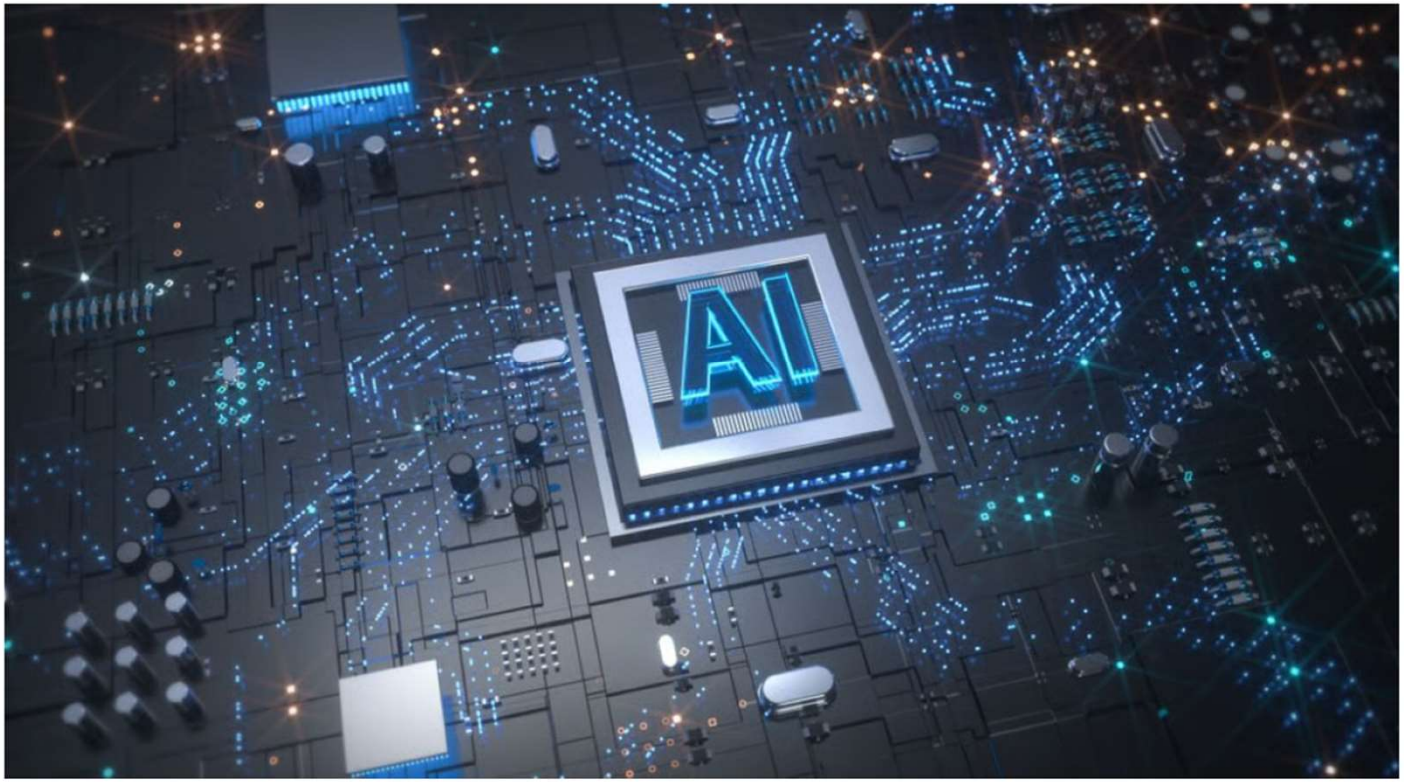
MediaTek and NVIDIA expected to unveil their Arm-based AI processor in the first half of 2025, will

By Mayank Sharma

April 23, 2024

Google and Meta unveil their Nvidia challengers

Custom AI chips can give Nvidia headaches



Meta and Google have just shown the way to take on Nvidia in the AI chips race

Meta and Google just announced their intention to push further into the AI chip race, currently dominated by Nvidia. Both companies are touting their new custom AI chips as a key element of their AI strategies.

6-Core CPU, 40-

Publ

Apple's M3 Max chip is expected to power the 14" and 16" inch MacBook Pro models in 202

IEEE Spectrum / Expect a Wave of Wafer

NEWS | SEMICONDUCTORS

Expect a Wave of Wafe
allows for one version no
in 2027

'Feels like magic!': Groq's ultrafast LPU could well be the first LLM-native processor — and its latest demo may well convince Nvidia and AMD to get out their checkbooks

News By Wayne Williams published 28 February 2024

onal bottlenecks

ate commission. Here's how it

NEWS FEED

Google's Tensor G5 processor to enter tape-out stage, manufactured with TSMC's 3nm process

by TechNode Feed Jul 2, 2024

News

NVIDIA plans ARM chip for laptops: A potential Apple M4 killer?

Philipp Briel - 13. August 2024

TSMC's wafer-scale integration tech is the key to

TESLA

Popular Posts

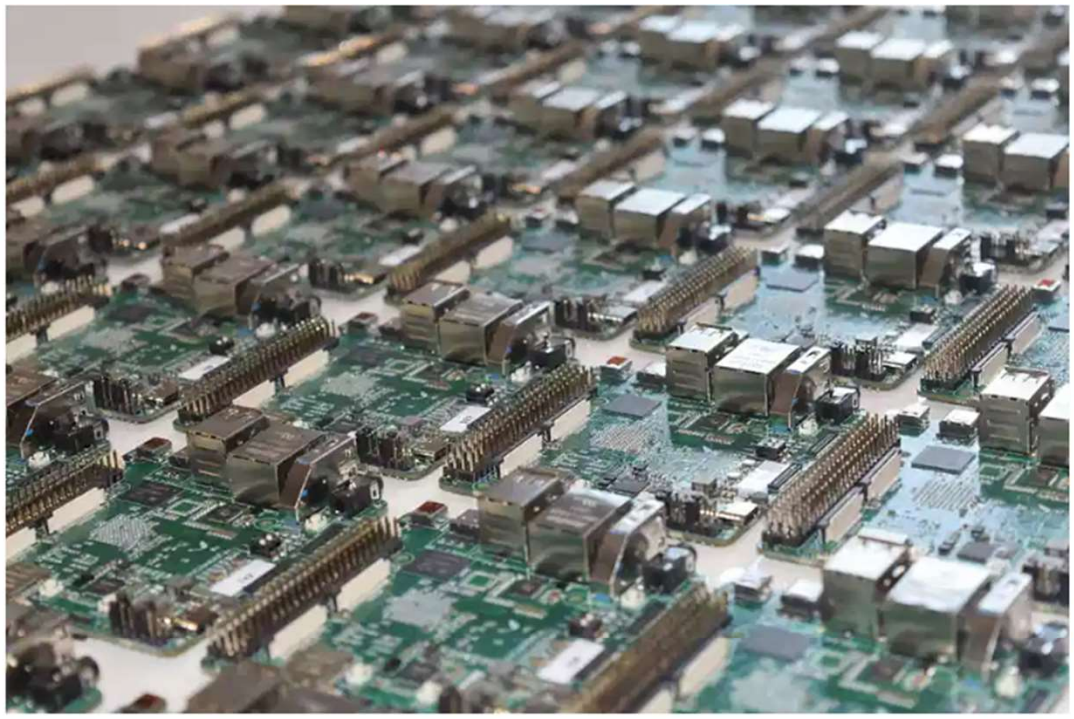
- Backforce V test: New refer in the 400 euro segment
- Backforce One Plus review: the competition look old aga
- Backforce One (Plus) Disco Coupon Codes
- Backforce One review: Can convince with a gaming cha
- Backforce versus noblechai thrones for a hallelujah
- Backforce versus Secretlab versus e-sports experience
- Coupons & Discount Codes

TECHNOLOGY

Innovation At The 'Edge': Hardware Startup From Surat Pioneers The Reconfigurable Edge Computer — And Reshapes Computing As We Know It

KARAN KAMBLE

May 25, 2024, 02:02 PM | Updated 02:02 PM IST



Vicharak's Vaaman edge computing board

Last month, Surat-based hardware startup Vicharak launched their very first product — an edge computing board called Vaaman.

News ar

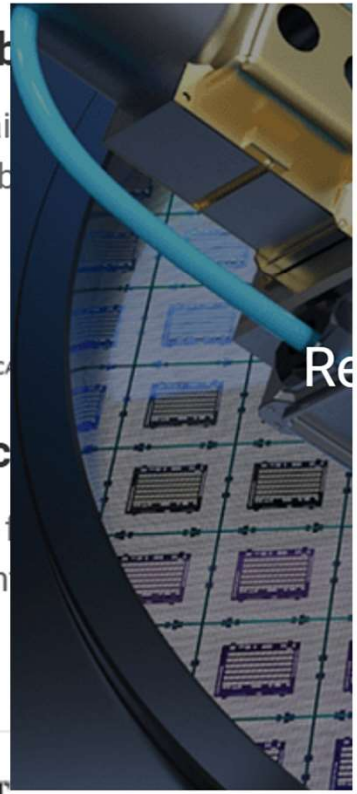
Indiatim TESSOLVE A HERO ELECTRONIX VENTURE

olve Labs Company Insight Co

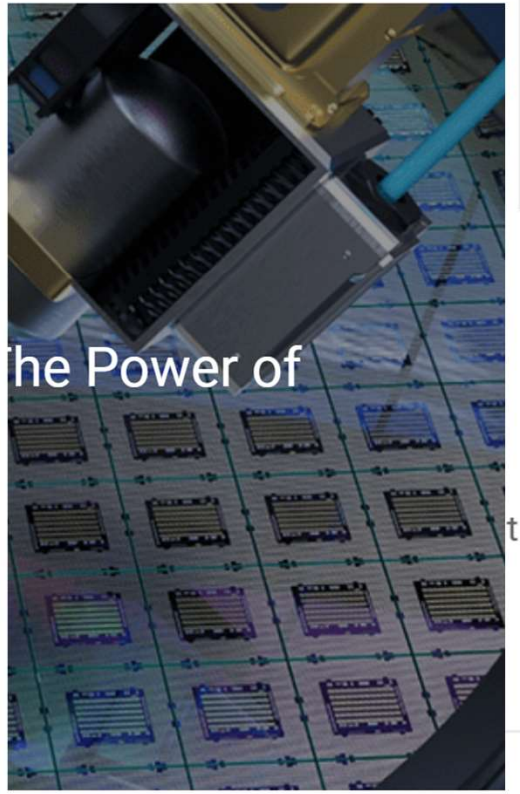
India b

Chennai

weight b



Re



he Power of

ars TECHNICA

Qualc

Arm is t t

the cen

THE T

Expl

Indic

parti

eled by strategic

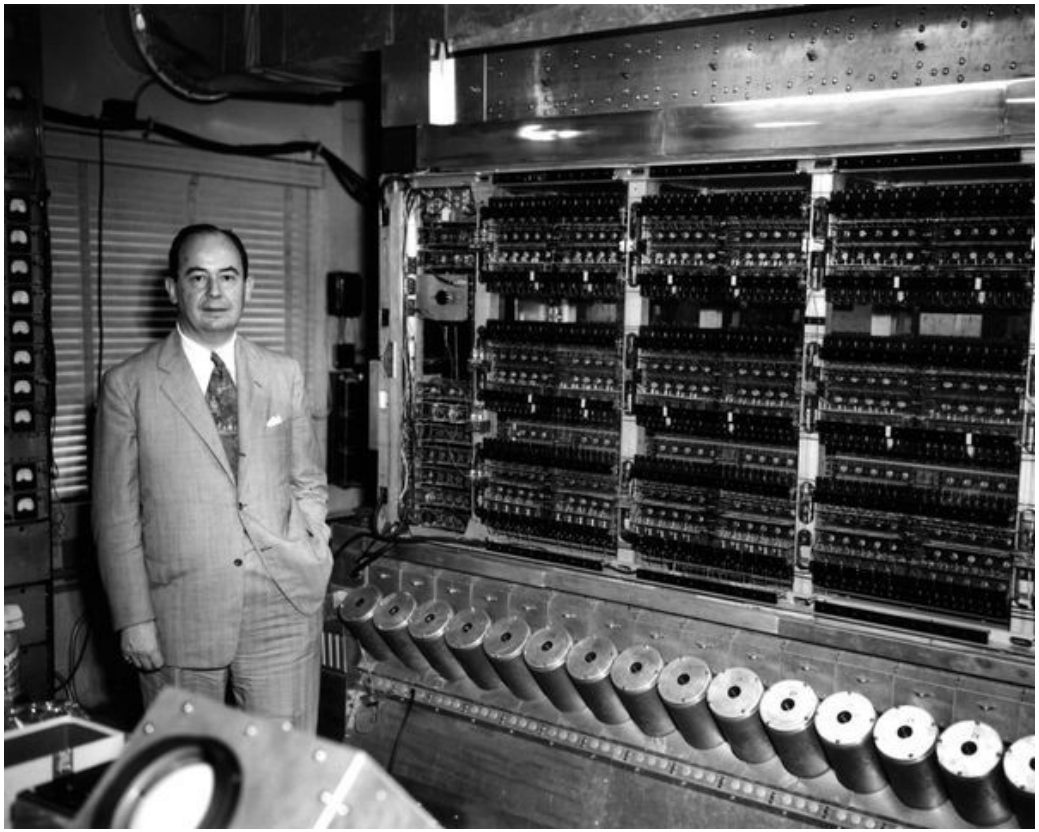
or (DIR-V) Programme.

India is promoting the use o conductor design, instead of

ARM technology that is ...

AARFIVE DESIGNS

Computers Then..



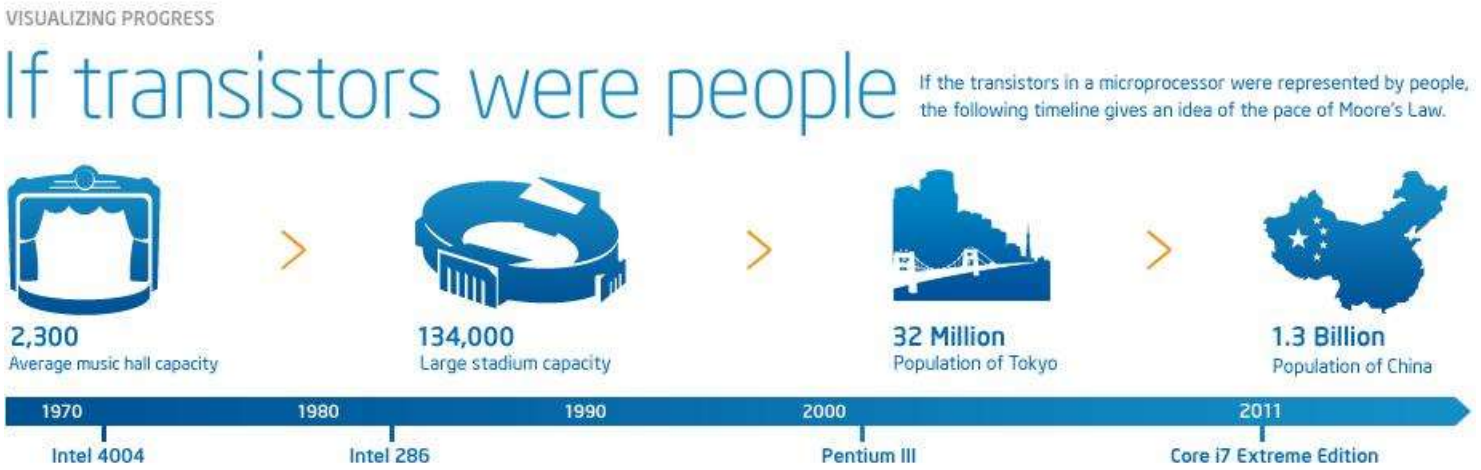
John von Neumann with the IAS Computer (Courtesy of the Shelby White and Leon Levy Archives Center, Institute for Advanced Study (IAS))

AARFIVE DESIGNS

Computers Now

AARFIVE DESIGNS

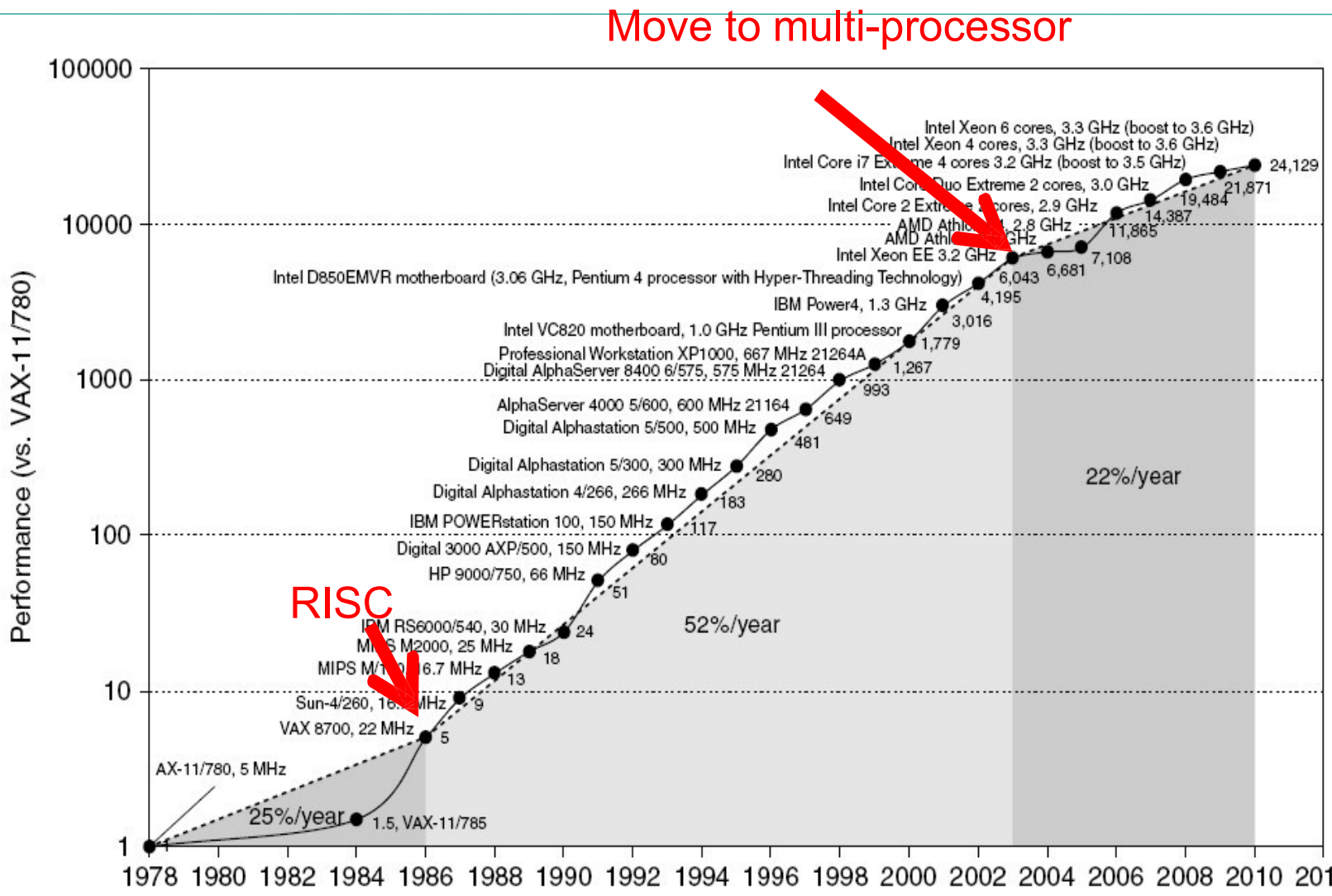
Moore's Law



Courtesy:

http://www.intel.com/content/www/us/en/silicon-innovations/moores-law-technology.html

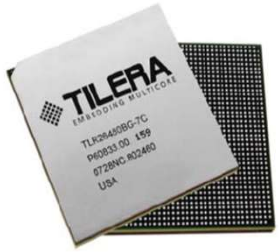
Sequential Processor Performance

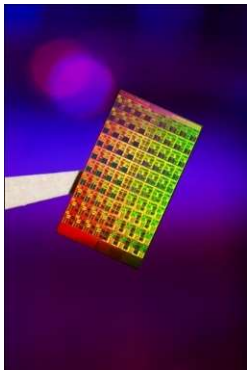


The era of Many-Core systems

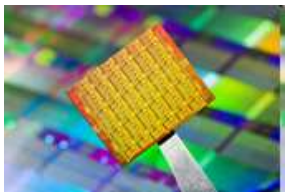
- How to keep up with demands on computational power?
  - Can not scale clock frequency
  - Solution: Increase number of cores - parallelism
    - Mass Market production of Intel, AMD dual-core and quad-core CPUs
    - Custom Systems-on-Chip (SoCs)
  - Many Core chips from Tilera for networking, cloud computing and multimedia applications.

Adapteva's Epiphany





Intel 80 core processor



Single-chip Cloud Computer

'Number of cores will double every 18 months'

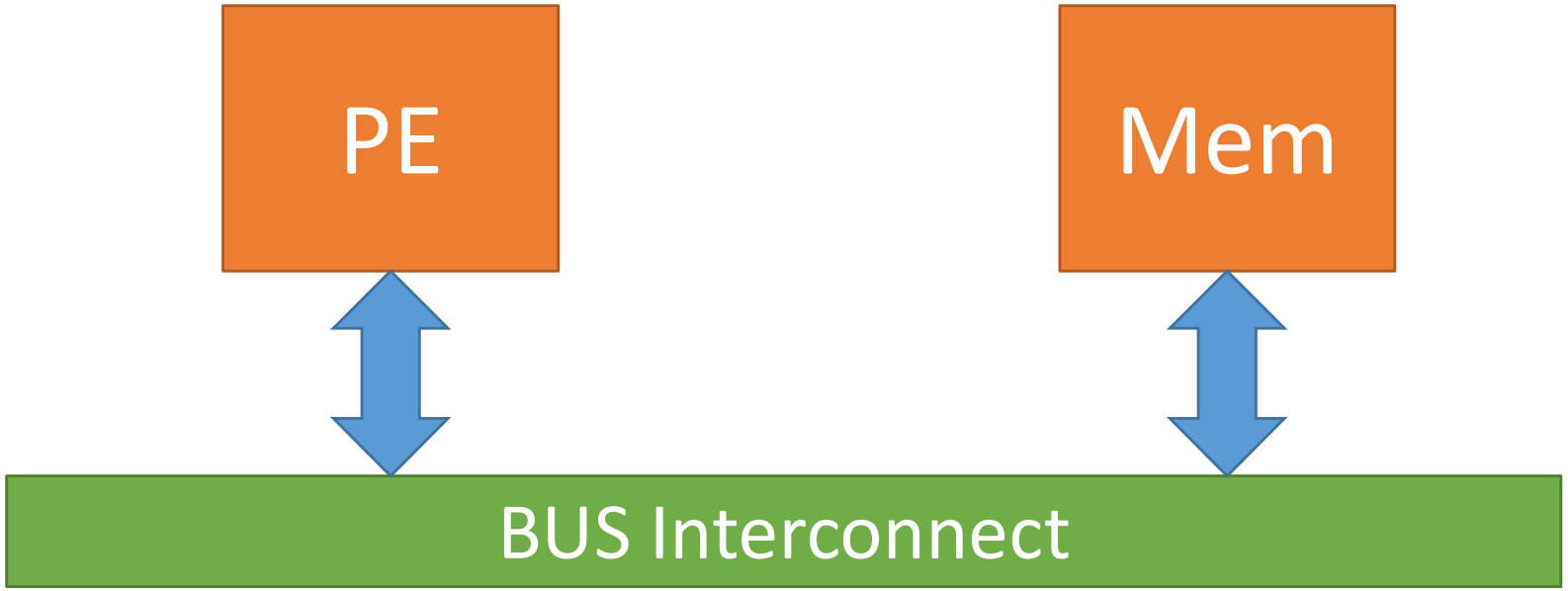
- Prof. A. Agarwal, MIT, founder of Tilera Corporation

AARFIVE DESIGNS

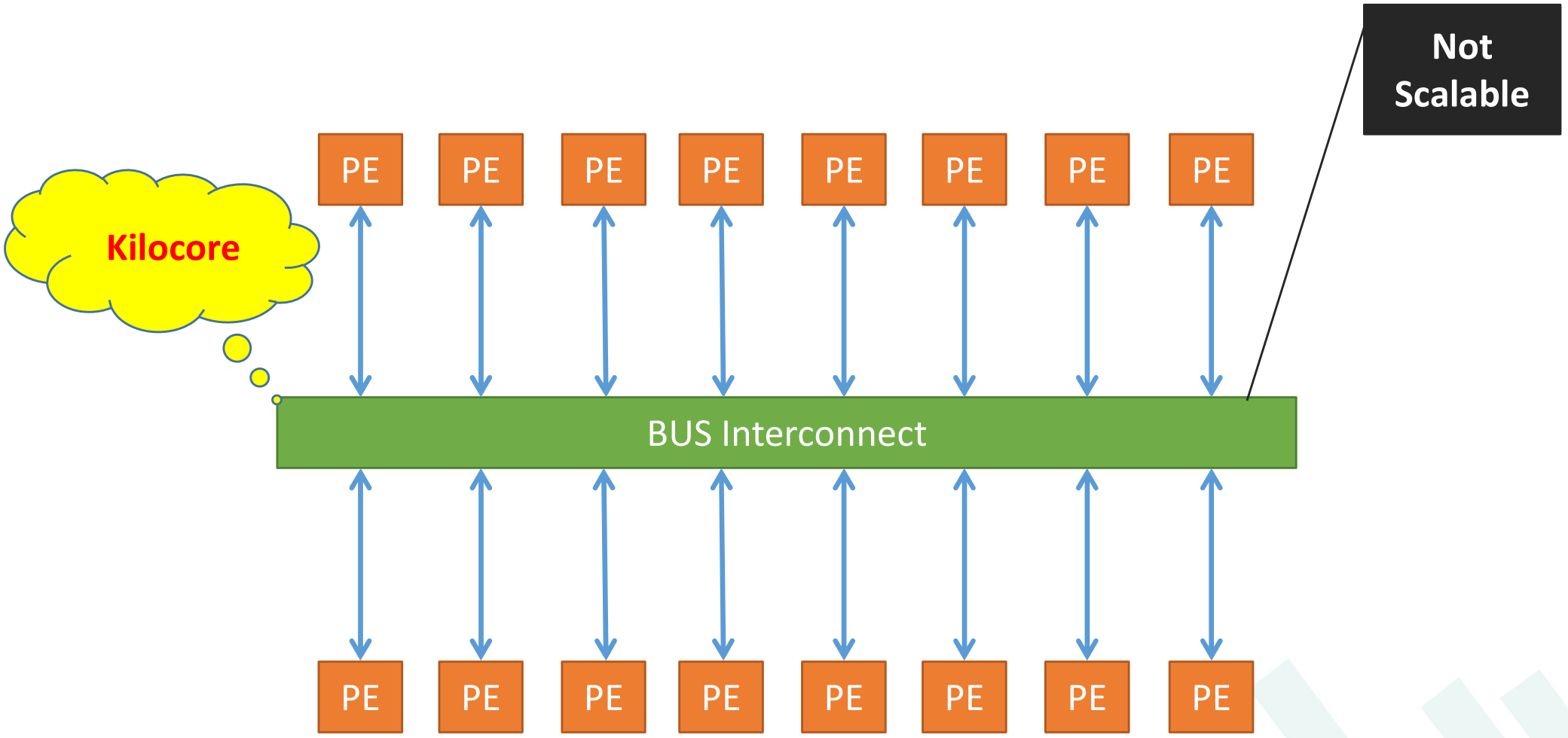
The era of Many-Core systems

- We are at the early stage of Many-core Processor evolution
  - Many-core is going to be ubiquitous
  - Immense possibilities:
    - Server-type performance on handheld devices

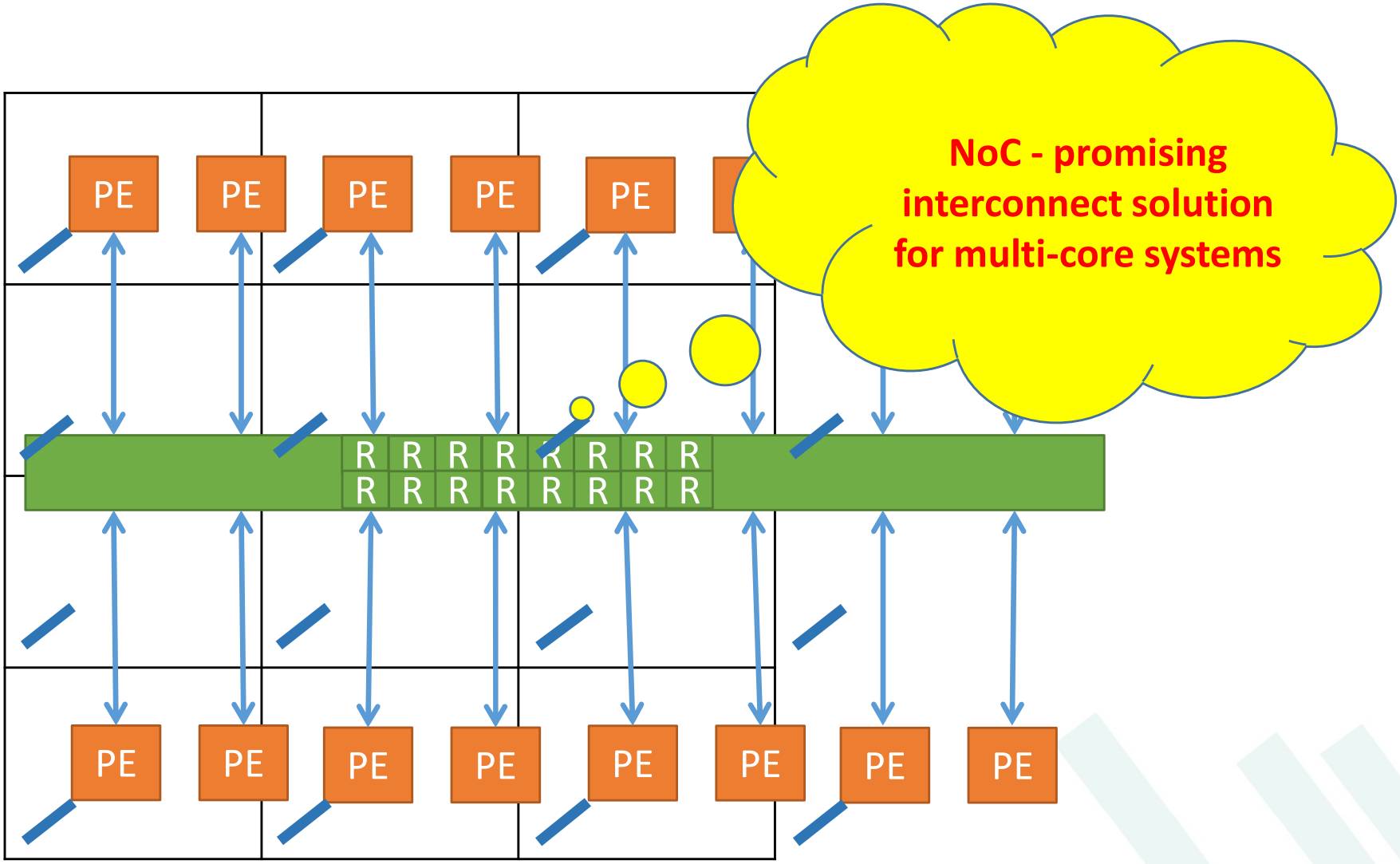
Basic Computing System



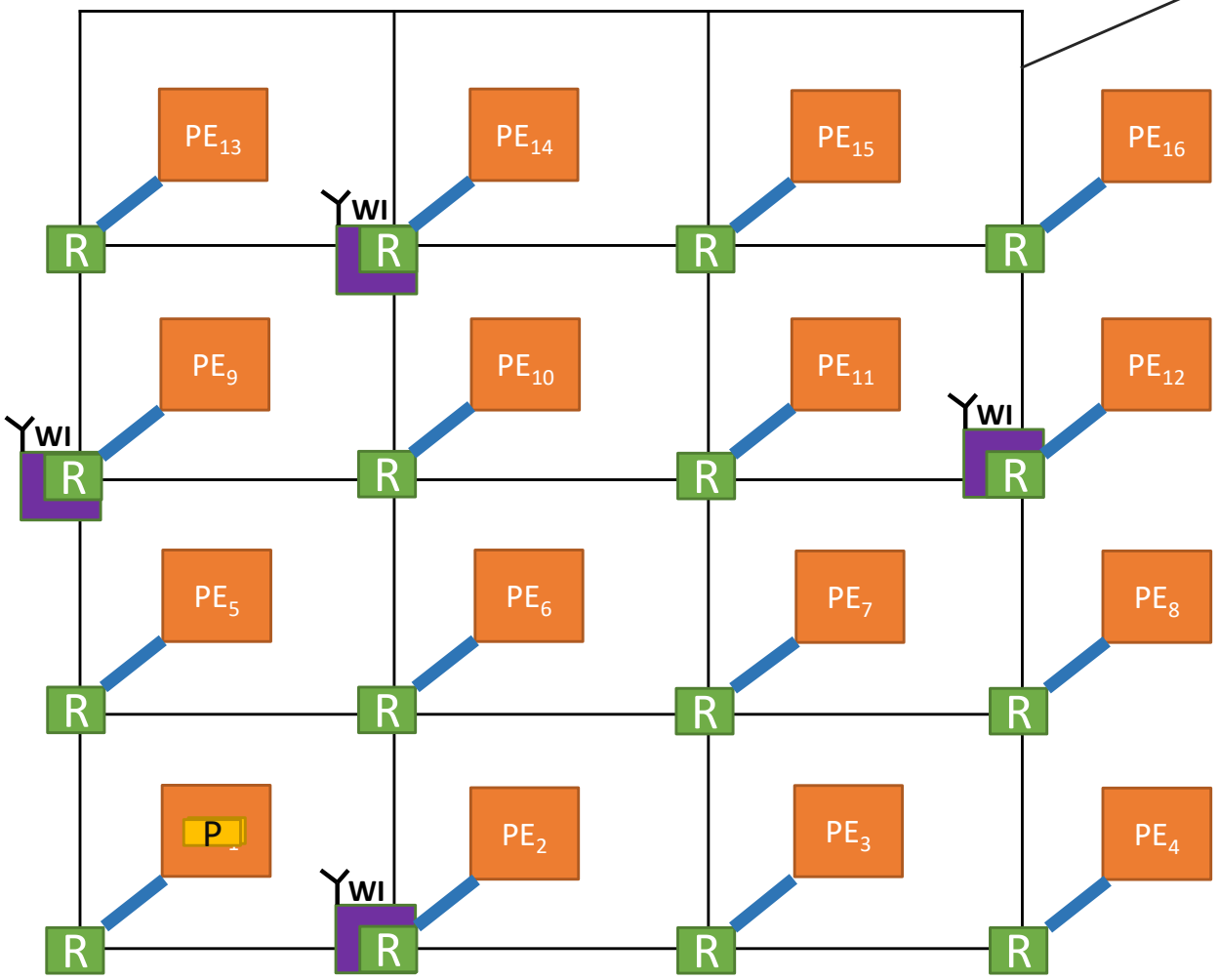
Multi-Core Systems



Network-on-Chip (NoC)



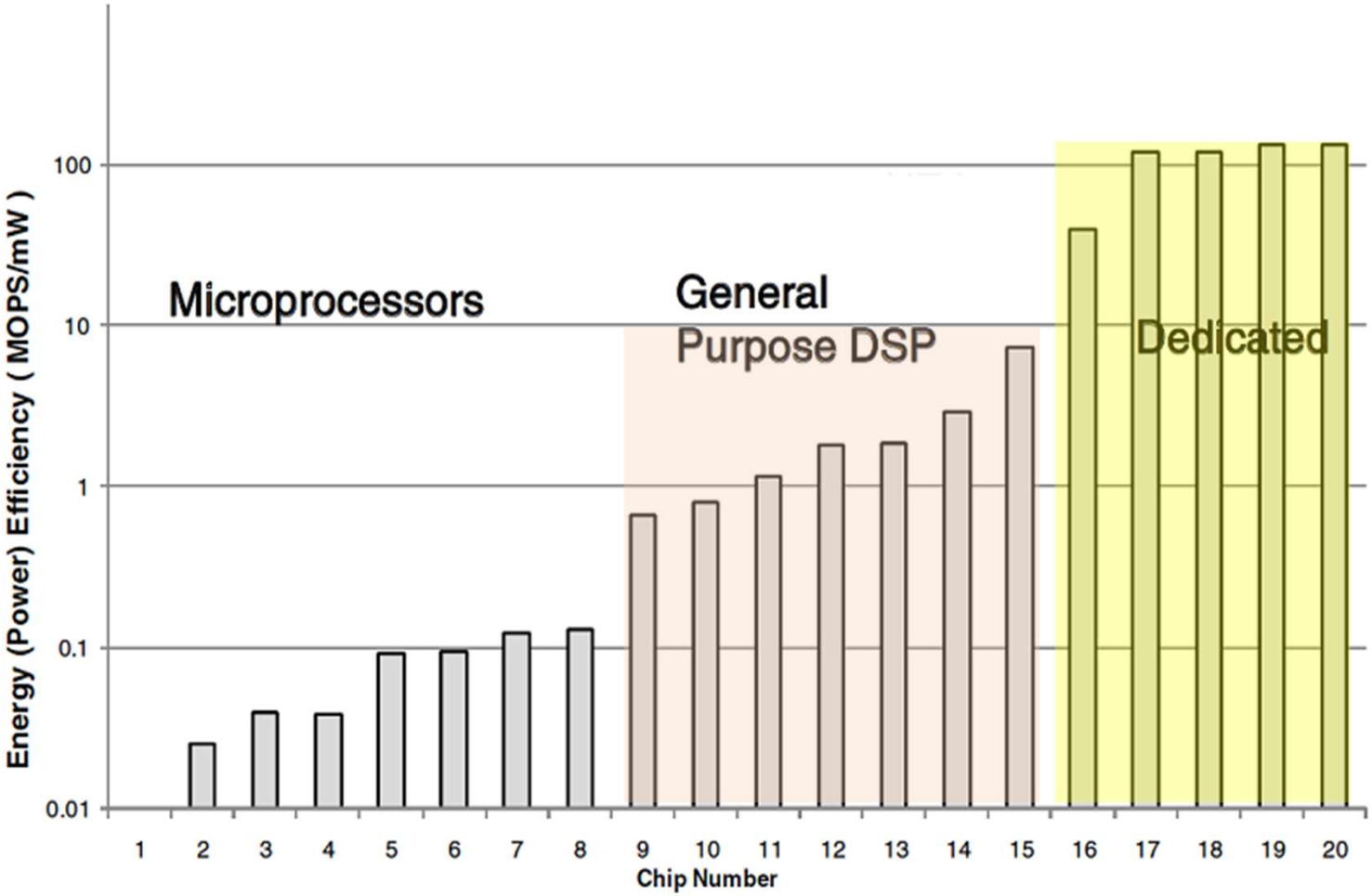
Wireless NoC (WNoC)



- Multi hop communication
- Power and performance bottleneck

AARFIVE DESIGNS

Efficient System Design Strategy



16	Encryption
17	Hearing Aid
18	FIR for disk read
19	MPEG Encoder
20	802.11 Baseband

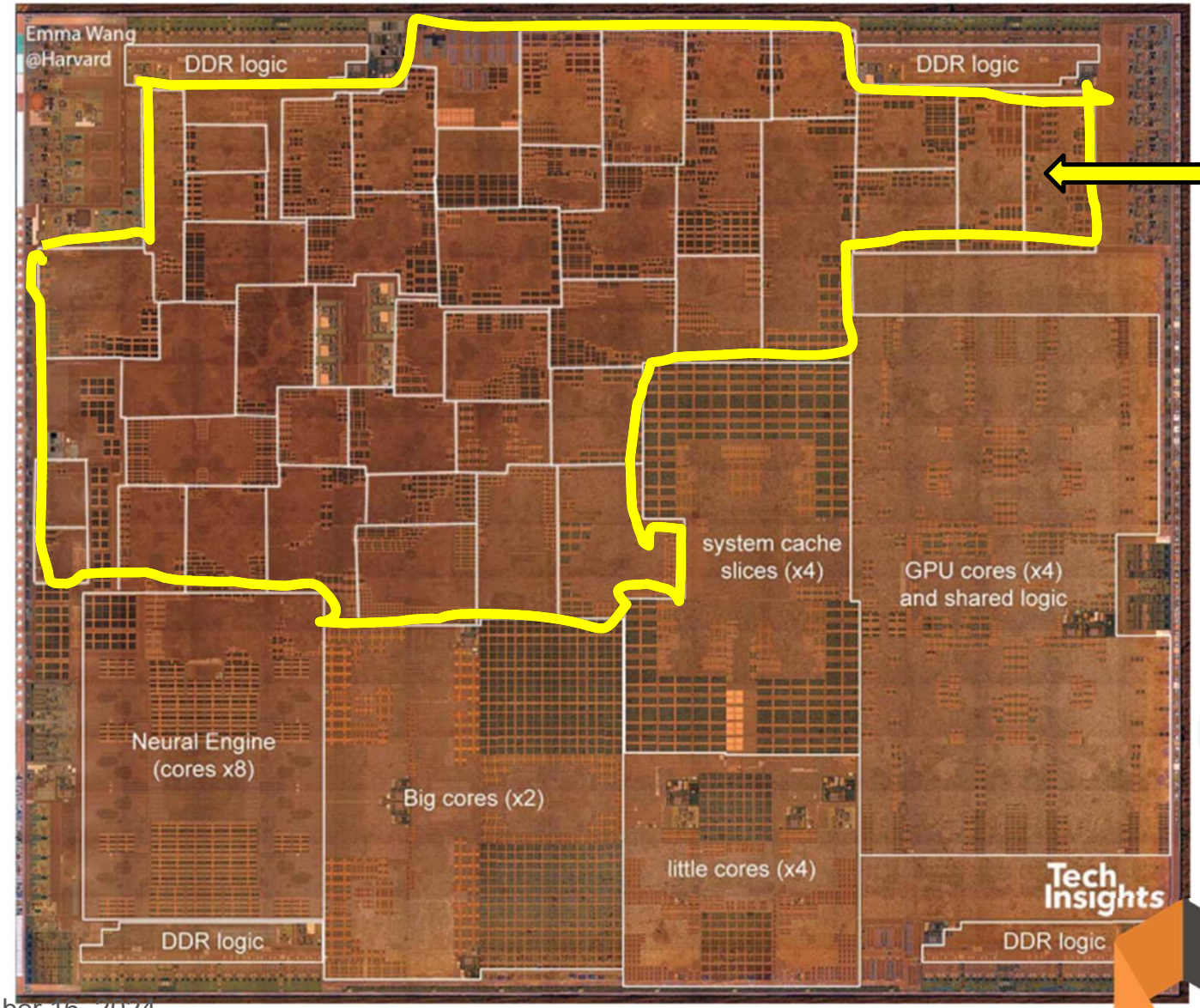
Fig. 1. Efficiency of specialized accelerators

[Zhang and Brodersen]

AARFIVE DESIGNS

Efficient System Design Strategy

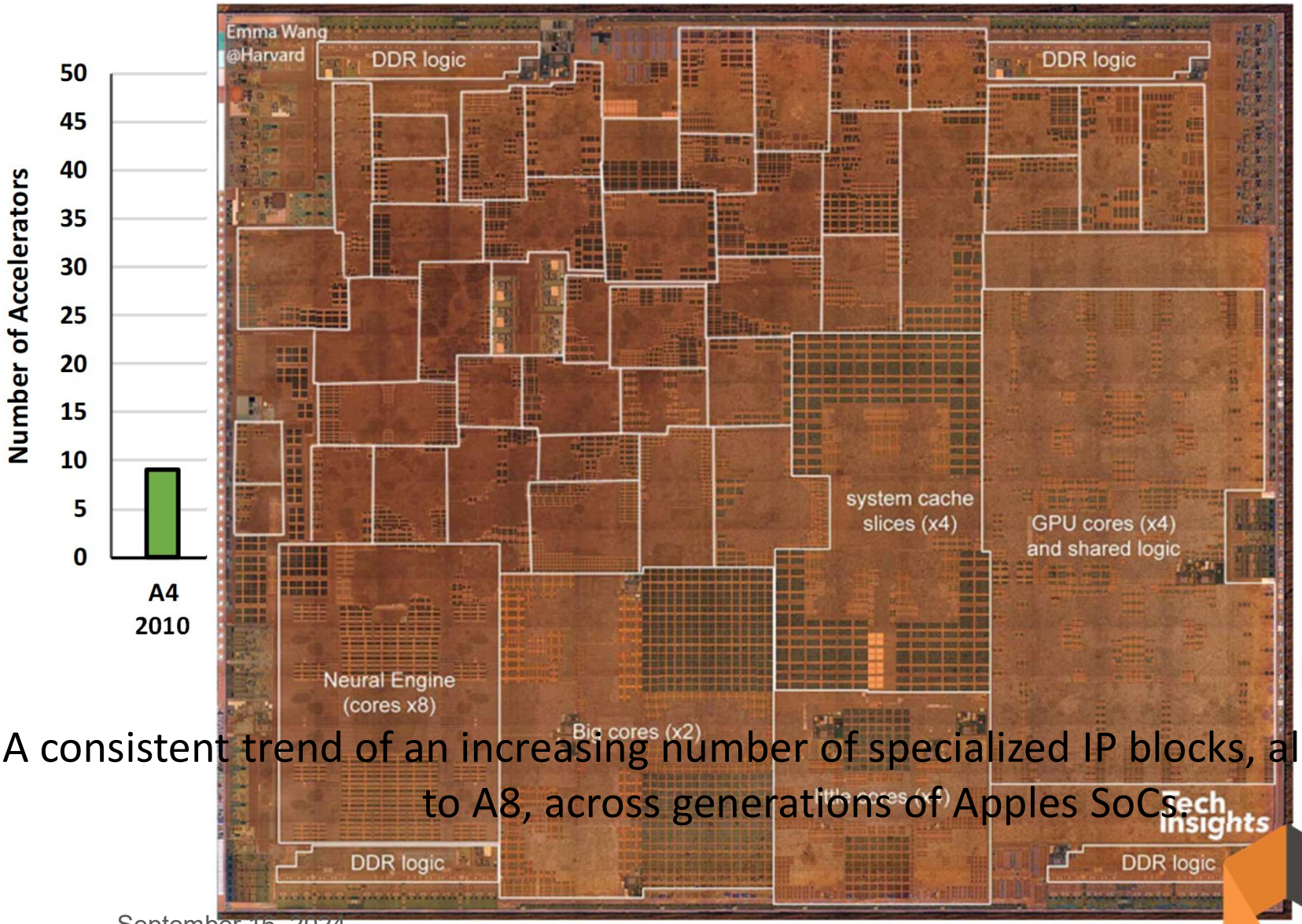
Apple System-on-Chip : A12



[Pic source: http://vlsiarch.eecs.harvard.edu/research/accelerators/die-photo-analysis/]

AARFIVE DESIGNS

Efficient System Design Strategy

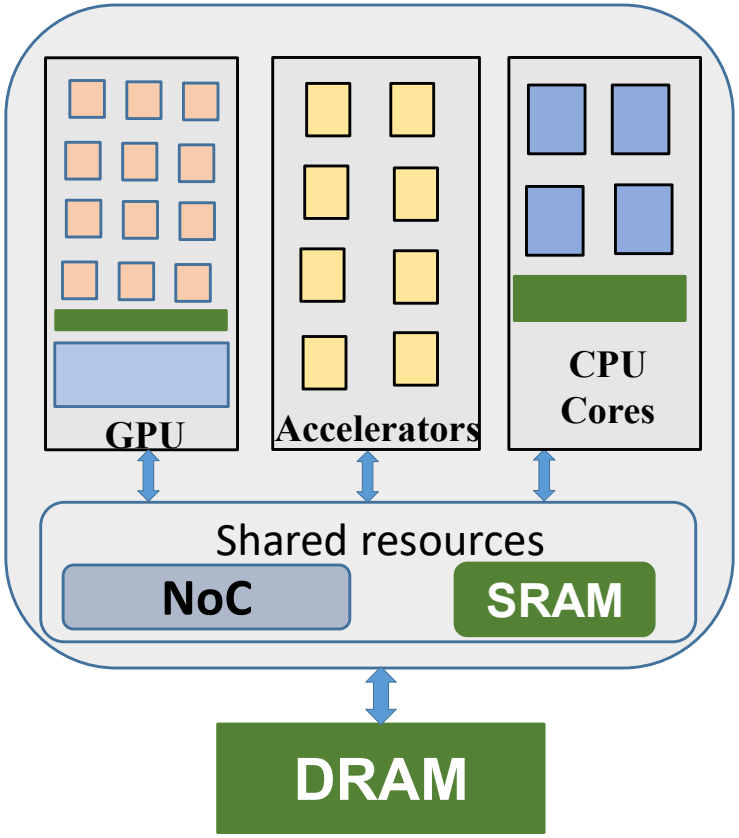


A consistent trend of an increasing number of specialized IP blocks, almost 3x from A4 to A8, across generations of Apples SoCs.

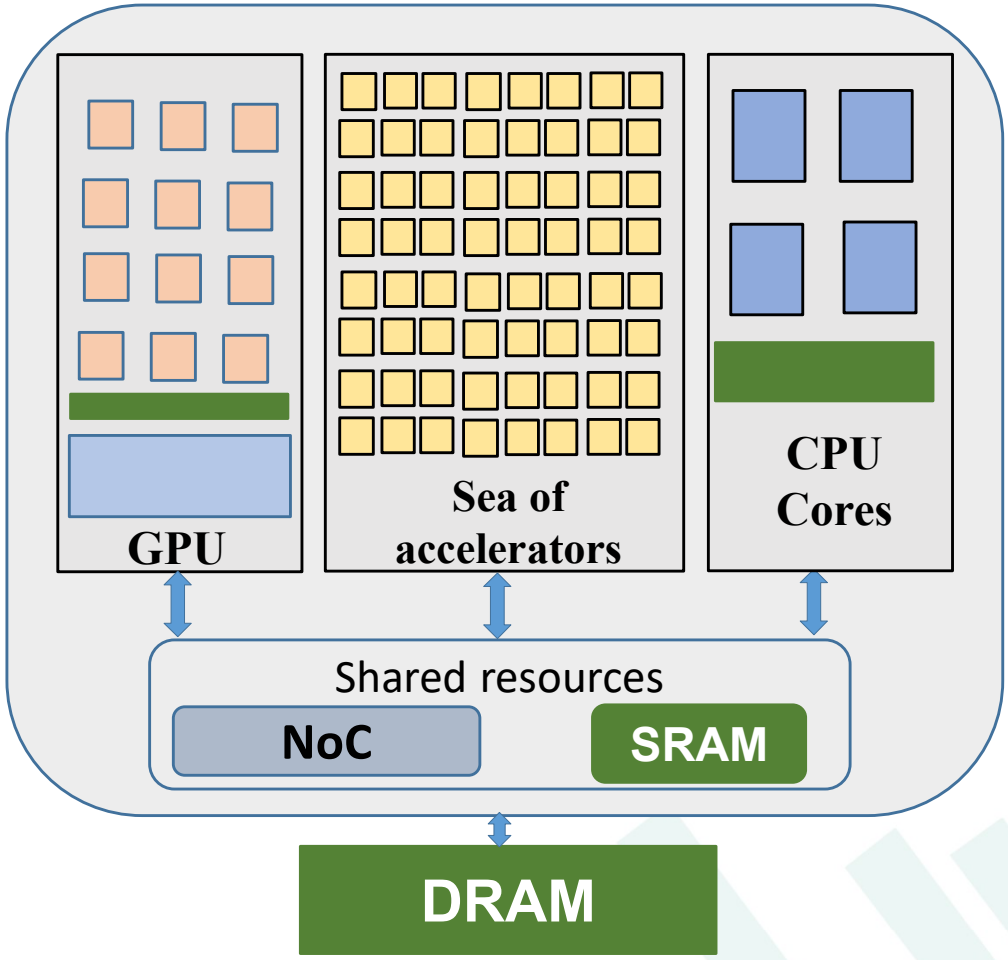
[Pic and data source: http://vlsiarch.eecs.harvard.edu/research/accelerators/die-photo-analysis/]

AARFIVE DESIGNS

Future of SoC



Present accelerator-rich SoC



Emerging accelerator-rich SoC

About AarFIVE Design Pvt Ltd

AARFIVE DESIGNS

About AarFIVE Design Pvt Ltd

People who **trust** us

Academy *Partners*

IP *Partner*

AARFIVE DESIGNS

Thank You!

AARFIVE DESIGNS

Contact Us

Website
www.aarfive.in

Email
business@aarfive.in

Address
IIIT - Delhi, near Govind Puri Metro Station, Shyam Nagar, Okhla Industrial Estate, New Delhi, Delhi, India

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

September 17, 2024